**Timetable and Programme for the CESSDA PPP Expert Workshop on
Harmonisation Issues in Comparative Social Surveys
Thursday 3 April 2008, Hotel Abba Montparnasse, Paris**

Each presentation should last around 15-20 minutes, to allow at least 10 minutes for discussion.

Room 1: "Metropolitain-Bolshoi", Room 2: "Scala".

Room 1		**Session 1**
09:00	Markus Quandt	Welcome and Introduction
09:15	Silke Schneider	"The International Standard Classification of Education (ISCED-97): Application to national educational qualifications and implementation in cross-national surveys"
09:45	Harry Ganzeboom	"Harmonizing Education and Occupation in Cross-National Comparative Research"
10:15	Gara Rojas-Gonzales	"Harmonizing microdata in EU-SILC"
10:45		*Coffee Break*
Room 1		**Session 2 A**
11:00	Peter Elias	"ISCO 88 and its pending revision"
11:30	Jonas Edlund	"Constructing nominal categorical class schemas: what kind of information is needed?"
12:00	Laurence Coutrot, (Annick Kieffer)	"ESeC: origins, major concepts, and implementation"
Room 2		**Session 2 B**
11:00	Willem Saris	"Measurement requirements for comparative research (with multiple reflective indicators)"
11:30	Christof Wolf	"Measuring religiosity and harmonizing religious measures in cross-national surveys"
12:00	Richard Topf	"The CCESD-IS (Centre for Comparative European Survey Data Information System)"
12:30		*Lunch break*
Room 1		**Session 3**
13:45	Claude Grasland	"Social data and the modifiable area unit problem"
14:15	Alexandre Kych	"Geographical location information and social survey data (working title)"
14:45	Ruud Luijkx	"COMPSOC: exploiting, documenting, and enriching COMParative data from large-scale surveys in the SOCial sciences"
15:15	Paul Lambert	"Distributing occupational information resources for comparative research: Experiences of the GEODE project"
15:45		*Coffee Break*
Room 1		**"Round Table"**
		Discussion on creating an harmonisation infrastructure
16:00	Markus Quandt	"Kick-off presentation: The approach of the CESSDA-PPP harmonisation work package"
	Ruud Luijkx, Joachim Wackerow, Christof Wolf	*Invited discussants*
17:15		*Approximate end*

The International Standard Classification of Education (ISCED-97):
Application to national educational qualifications and implementation in cross-national surveys

Silke L. Schneider

This presentation will summarise and synthesize the results of more than two years' work of an international team of researchers evaluating the application of the ISCED-97 to national educational qualifications in 15 European countries. This collaboration took place in the context of the EQUALSOC network (research team "Evaluation of the ISCED-97 for Comparative Research") and resulted in a book to be published by MZES (Mannheim).[*]

The core problems in the application of the ISCED-97 can be divided into three areas:

1. conceptual limitations of the ISCED-97 itself (heterogeneity of specific ISCED sub-categories),

2. constricted implementation in cross-national surveys (usually involving a simplification of the ISCED-97), and

3. difficulties in the application of the ISCED-97 to actual educational qualifications in specific countries (e.g. due to educational reforms and changing educational qualifications or shortcomings of national data collection procedures).

The consequences of the first two issues for cross-national research will be illustrated with an example. Finally, some suggestions for refining the measurement of educational attainment in cross-national surveys will be made. These involve 1) the refinement of the ISCED-97 in a future revision, 2) the amelioration of recoding procedures for educational attainment variables in cross-national surveys and 3) the enhancement of data collection procedures and thus nationally specific measures of educational attainment.

Contact: Silke Schneider, Nuffield College, New Road Oxford OX1 1NF
E-mail: Silke.schneider@nuffield.ox.ac.uk

Harmonizing Education and Occupation in Cross-National Comparative Research

Harry B.G. Ganzeboom

In this presentation I compare procedures for harmonizing occupation and education in cross-national comparative research. At first glance these two problems have great similarity, and they have often been approached with the same principles and procedures. I argue that in fact the two problems have quite different answers. The best way to harmonize occupations is to use a similar question and answer format and to code or classify the answers in the International Labour Office's International Standard Classification of Occupations [ISCO]. Measurement quality can be improved by asking parallel questions on occupation. Multi-trait multi-methods [MTMM] models using multiple measures show that very little information is lost when moving from local to cross-national measures to scale occupation and that nationally sensitive coding is to a large extent superfluous. Both cross-national and national research would be served well when ISCO was employed throughout; even historical comparisons can be validly conducted with modern classifications (and vice versa: comparison of contemporary data works well with older classifications).

Occupations and the underlying division of labour are produced by technological and economic structures that are to a large extent universal to all societies. This makes for the high degree of similarity and

---

comparability between societies, even in the long historical range. Education, by contrast, is totally man-made and institutionally determined. The immediate consequence of this is that educational structures and degrees are hard to compare, both cross-nationally and historically. Comparability is in fact obstructed when one uses a cross-national classification, in particular Unesco's International Standard Classification of Education [ISCED] to measure education in an international comparative scale. Unlike ISCO, with its great detail and historical pedigree, ISCED is not a detailed and universal standard classification at all, but rather a crude aggregation of educational degrees into a small number of categories for a small number of countries at a given point in time (1997). One can only expect that such a measure glosses over important distinctions in local educational structures and is a rather hopeless instrument to compare over time. This last feature makes it next to useless in general population research, in which we find a mixture of educational cohorts.

I argue that the best way to measure education is a locally and historically sensitive measure, like used in ISSP and ESS (for respondents). For many practical problems satisfactory post-harmonization can be achieved by using within-country standardized scores, such as percentiles. This strategy exploits two of the most important characteristics of any educational measure, namely (A) the fact that educational degrees always and everywhere have a strongly ordinal character, (B) that the optimal order is easy to establish because in each context only a fairly limited number of distinctions is used. I continue to outline a methodology that would go one step further in educational measurement, which is using anchor points to establish a common educational metric for all countries at all times. By using multiple indicators (using duration as a parallel measure) the loss of information and bias can be assessed in a MTMM model.

Contact: Prof. Dr. Harry Ganzeboom, Free University Amsterdam, Faculteit der Sociale Wetenschappen - Methoden & Technieken, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands
E-mail: HBG.Ganzeboom@fsw.vu.nl

Harmonizing microdata in EU-SILC

Gara Rojas-Gonzales

The EU Statistics on Income and Living Conditions (EU-SILC) is the EU reference source for comparative statistics on income distribution and social exclusion at European level. This paper analyses the effort from both Eurostat and European countries towards a high degree of comparability of data from this survey.

Countries provide two types of annual data: cross-sectional and longitudinal data; and this data are used to compile several EU indicators on social inclusion, pensions and health and long term care. They also provide quality reports analysing the accuracy, coherence and comparability of their data.

Eurostat describes quality of data with six components: relevance, accuracy, timeliness and punctuality, accessibility and clarity, coherence and comparability. Since its setting up, quality has been a major concern for EU-SILC, and especially comparability, which is studied in this paper. Comparability is not a one shot process; it requires continuous monitoring and adjustment.

In contrast with ECHP (European Community Household Panel), its predecessor, EU-SILC does not have an input harmonized structure but an ex-ante output harmonization strategy. As a matter of fact, ECHP was based on the idea of a common survey and EU-SILC is based on a common framework. This common framework is defined by harmonized lists of target primary and secondary variables, common concepts, a recommended design, common requirements (for imputation, weighting, sampling errors calculation) and classifications aiming at maximising comparability of the information produced. Discussions on the variables and concepts to include in the new survey replacing ECHP began in 1999-2000 focusing on the outcomes, and before EU-SILC became active in 2004 all the concepts were already established. This means that the framework is flexible in terms of data sources and sampling design as long as the indicators provided fulfil certain characteristics and could be consequently comparable.

This paper reviews the organization in EU-SILC presenting Eurostat general strategy with respect to harmonization and tries to assess the outcome in terms of comparability.

Contact: Gara Rojas-Gonzales, Eurostat - Unit F3 - Living conditions and social protection statistics, Office: BECH B2/432

E-mail: Gara.Rojas-Gonzales@ec.europa.eu

ISCO 88 and its pending revision

Peter Elias

The International Standard Classification of Occupations (ISCO) underwent a major overhaul in the late eighties. ISCO'88 made explicit the concepts of 'skill levels' and 'skill specialisations' to create a classification that had a clearer conceptual basis than had hitherto been the case. Despite the efforts that went into this work, many Western European countries were slow to adopt the new standard, often choosing to create a 'crosswalk' between their national classification and ISCO'88. For the countries of Eastern Europe, the need for a new standard classification was more urgent. The realignment of their national statistical systems with EU requirements meant that ISCO'88 (or the EU variant of the standard) was introduced as their national classification. As a result, we have significant variations across Europe in terms of the national experience of using ISCO and the comparability of the resulting occupational statistics.

Twenty years on Eurostat has the opportunity to review this situation with the introduction of ISCO'08. This presentation focuses on the differences between ISCO'88 and ISCO'08, examining in particular the areas where interpretation of the standard has been problematic and highlights new problem areas. The presentation concludes with information about the plans that Eurostat has for the introduction of the new standard.

Contact: Prof. Dr. Peter Elias, University of Warwick, Warwick Institute for Employment Research, Social Studies Building, Coventry, CV4 7AL, UK

E-mail: Peter.Elias@warwick.ac.uk

Constructing nominal categorical class schemas: what kind of information is needed?

Jonas Edlund

The presentation discusses a selected number of nominal categorical class schemas (e.g., the Erikson Goldthorpe classification of occupations; Ousch's classification of occupations; Grusky's micro-class schema.

1. A brief description of the theories underlying the class schemas.

2. In order to make a reliable class classification, what kind of variables are important to include in a survey?

3. Looking forward: class classification and harmonization over time and across surveys

Contact: Ph.D. Jonas Edlund; UMEÅ University, Sociologiska institutionen, SE-901 87 Sweden

E-mail: Jonas.Edlund@soc.umu.se

EseC: origins, major concepts, and implementations

Laurence Coutrot

Abstract not available

Contact: Dr. Laurence Coutrot, Centre Maurice Halbwachs, Centre National de la Récherche Scientifique, École Normale Supérieure, 48 Boulevard Jourdan, 75014 Paris
E-mail: Laurence.Coutrot@ens.fr

Measurement requirements for comparative research (with multiple reflective indicators)

Willem E. Saris

The measurement requirements for comparative research specified in the literature for concepts with reflective indicators are: configural invariance, metric invariance and scalar invariance. Configural invariance is obtained if the same standard factor analysis model should hold for all different groups. One has to add to this requirement the equality of the loadings to speak of metric invariance and to obtain scalar invariance the intercepts should also be equal in the different groups.

In this presentation we argue that these requirements are too strict. We will specify a response model that makes a distinction between a measurement part of the response process and the cognitive part. We will argue that it can happen that the above specified requirements are not satisfied because of differences in the measurement process or in the cognitive part or in both. It is essential that the cognitive part is the same across groups because otherwise people have different ideas about the concepts of interest. Differences in the measurement part are less fundamental. One can estimate the differences in this measurement process separately and correct for these differences.

As a consequence we suggest that the above mentioned requirements for comparative research should hold after correction for measurement errors.

Contact: Prof. Dr. Willem E. Saris, ESADE, Department of Quantitative Methods, Universitat Ramon Llull, Av. de Pedralbes 60-62, E-08034 Barcelona
E-mail: Willem.saris@esade.edu

Measuring religiosity and harmonizing religious measures in cross-national surveys

Christof Wolf

The presentation consists of three parts. First, three established approaches to conceptualize and measure religiosity are presented:

- The dimensional approach originating in the work of Charles Glock;

- the concepts of intrinsic and extrinsic religiosity developed by Gordon Allport;

- a syntheses of these two aforementioned approaches proposed by Stefan Huber.

In the second part I will give an overview of the measures of religiosity typically found in cross-national surveys; these measures will be organized according to the theoretical framework laid out previously. Thereby it will become clear where the blind spots of our current data are.

In the next part I will draw attention to the different ways central measures of religiosity are operationalised and will show how they can be harmonized to yield valid cross-national comparisons.

In the final part of this presentation I will present some recommendations for the measurement and harmonizing of data on religiosity in cross-national surveys.

Contact: PD Dr. Christof Wolf, GESIS-Zuma, Postfach 12 21 55, 68072 Mannheim
E-mail: christof.wolf@gesis.org

The CCESD-IS (Centre for Comparative European Survey Data Information System)

Richard Topf

The ESF's *Beliefs in Government (BIG)* project first drew the attention of the international social science community to the lack of suitable research tools in CESSDA Archives to facilitate the harmonisation of survey data cross-nationally, over time. The CCESD-IS project was initiated as proof-of-concept research to address such problems through a radical reconceptualisation of the processes of meta-data and data storage and retrieval.

The key premises of the design concept include:

- to use the individual, full-text, survey *question*, rather than the complete *survey*, as the basic unit for archiving, manipulation, harmonisation, retrieval, and descriptive analysis

- at the same time, to ensure that original metadata and data are always accurately preserved, archived and fully retrievable

- to assume that the needs of the social science community for data processing are no more specialist nor complex than those of other public or commercial institutional sectors, and thus can best be met by using industry-standard, applications software

- to assume that, for practical purposes, off-the-shelf computer hardware and software has evolved to the point where it is unproblematic, simultaneously to pool and manipulate an infinite amount of metadata and data in 'virtual space'

- to build question-level, mapping of literally equivalent items into the system, but at the same time, to provide end-users with real-time, interactive tools to retrieve, codify and analyse choices of conceptually or functionally equivalent items on-the-fly

This presentation will include illustrations of the application of full-text, fuzzy-logic database tools for harmonisation of attitudinal survey items, such as political participation.

Contact: Prof. Dr. Richard Topf, Centre for Comparative European Survey Data Information System (CCESD-IS), Calcutta House, Old Castle Street, London E1 7NT
E-mail: topf@ccesd.ac.uk

Social data and the modifiable area unit problem

Claude Grasland

The talk will present results from research performed under the ESPON program. For complimentary information, please visit:

http://www.espon.eu/mmp/online/website/content/projects/261/431/index_EN.html

and, with particular relevance to this talk, pp. 50-57 of the final report of ESPON:
http://www.espon.eu/mmp/online/website/content/publications/98/1232/file_2480/scientific-reportii_web.pdf

Contact: Prof. Dr. Claude Grasland, Université Paris 7, UFR GHSS (Géographie, Histoire et Sciences de la Société), Case courrier 7001, 2, place Jussieu, 75251 Paris Cedex 05
E-mail: claude.grasland@parisgeo.cnrs.fr

Geographical location information and social survey data (working title)

Alexandre Kych

Abstract not available

Contact: Alexandre Kych, Centre National de la Récherche Scientifique
E-mail: alexandre.kych@ens.fr

COMPSOC: exploiting, documenting, and enriching COMParative data from large-scale surveys in the SOCial sciences

Ruud Luijkx

Abstract not available

Contact: Dr. Ruud Luijkx, Tilburg University, Faculty of Social and Behavioural Sciences, PO Box 90153, 5000 LE Tilburg, The Netherlands
E-mail: R.Luijkx@uvt.nl

Distributing occupational information resources for comparative research: Experiences of the GEODE project

Paul S. Lambert

Occupational information resources' provide information on occupations published by, and useful to, social researchers. They take the form of national and cross-national databases on occupations, made available at internet sites and through alternative formats. Many such databases exist, including translation tables and matrices with recommendations on where occupational units should be located within social classification schemes; and databases of descriptive information about occupational units, such as summary statistics on different occupations.

Simultaneously, most social science micro-data collections record some form of data about respondents' occupational positions. These could be fruitfully connected with numerous suitable databases on occupations. Nevertheless, non-specialist social scientists have typically not shown great fluency in connecting together their micro-data with suitable occupational information. On the contrary, it is easy to find examples of analysis of occupational data, particularly in cross-nationally comparative research, where data on occupational positions may appear to have been analysed in a sub-optimal manner. Better practice in exploiting appropriate occupational information resources could improve the ways in which social researchers understand their data on occupations and related concepts of social position, class and stratification, and would be particularly likely to make contributions to the understanding of occupational data in comparative research.

The GEODE project ('Grid Enabled Occupational Data Environment) is an ongoing endeavor to improve the quality and accessibility of occupational information resources for social science research. The project tries to cater both to the producers of occupational information resources (providing an online depository for newly published occupational information); and to those non-specialists who merely wish to connect their original data on occupations with suitable analytical information (the project provides an online 'portal' for searching for, and for matching together, suitable occupational information for relevant micro-data).

The GEODE project uses two main strategies to achieve this. The first is to draw up typologies for recording data about occupational information (meta-data). These allow diverse databases about occupations (relevant to different countries, time periods and other contexts, and released in a variety of

different formats) to be stored and made accessible in a consistent and navigable structure. The second strategy is to set up an internet based facility attempting to provide user-friendly navigation tools and instructions for depositing and accessing relevant occupational data.

At present, the GEODE project supports an internet service which can provide informative services to the producers and potential users of occupational information, but which, hitherto, has not yet been widely adopted by its potential users groups. In ongoing work, researchers on GEODE are reviewing usability of the GEODE services and seeking ways to improve the accessibility and subsequent uptake of the resource

Whilst working with diverse databases on occupations, the sociologists involved in the GEODE project were also increasingly engaged in methodological assessments of approaches to harmonization and comparability in occupational data. One topic which the GEODE research engages with is the extent to which harmonized analyses of occupational data could use 'specific' rather than 'universal' measures of occupational data (that is, code occupations using national- and time-specific schemes for the purposes of a pooled comparative analysis). It is asserted that provisions such as GEODE should help researchers achieve 'specific' approaches more easily. These may, ultimately, lead to significant improvements in the extent to which cross-national comparative analyses successfully engage with the local contexts which they review.

The way in which the GEODE services are provided exploits a computer science approach known variously as the 'Grid' and 'e-Science'. Both of these service provisions were developed with extensibility in mind, but were also specifically tailored around particular requirements for working with occupational data. The GEODE project has now been integrated into a wider project using this approach, called 'Data Management through e-Social Science' (www.dames.org.uk). In this presentation, the way in which the GEODE activities are to be extended to other specialist areas of information provision (educational data and information on ethnicity and immigration) will also be briefly introduced.

Contact: Dr. Paul Lambert, Stirling University, Department of Applied Social Science, Colin Bell Building, Stirling, FK9 4LA, Scotland, UK
E-mail: Paul.lambert@stirling.ac.uk